

Python for Data Analytics

Course Outline

Overview:

In this course, you'll learn the fundamentals of data analysis with Python. By the end of this course, you'll know how to read data from multiple sources like CSVs, SQL and Excel, and process that data using multi-dimensional arrays in NumPy, manipulate DataFrames in Pandas, and visualize it using Matplotlib and Seaborn, use web scraping packages to extract data from websites, gain insight into basic machine learning concepts and techniques and use GenAI tool on coding.

Course Duration:

4 days

Prerequisites:

There are no prerequisites but some programming experience is helpful.

Course Outline:

1. Python Review

- A Python Development Environment
 - Overview of Python scripting tools and IDEs: VS Code, PyCharm, Google Colaboratory (Cloud)
- A Review of Python language
 - Built-in types, data structures, mathematical and logical operations, control flow
 - Function, String, Math, Class, Date Structure
 - File Read/Write

2. NumPy Basics

- Features of ndarrays, basic methods and attributes
- Create, access, modify, and sort multidimensional arrays in NumPy
- Perform element-wise operations on ndarrays
- Use broadcasting to perform operations on ndarrays of different sizes

3. Panda Basics

- Features of Series and DataFrames, basic methods and attributes
- Create, access, and modify the main objects in Pandas Series and DataFrames
- Perform arithmetic operations on Series and DataFrames
- Load data into a DataFrame
- Indexing and Access Techniques

4. Pandas: DataFrame Basics

- DataFrame Construction
- DataFrame Change and Reorganization
- Indexing and Access Techniques
- Grouping, Pivoting, and Reshaping
- DataFrame CRUD operations - creating, reading, updating, and deleting data

Python for Data Analytics Course Outline

5. Data Visualization

- Visualization Is Storytelling
 - Colors, Format, Layout
 - Types of charts: Line plot, Scatter plot, Histogram, Box plot, Heatmap
- Matplotlib for Data Visualization
 - Steps for Creating a Data Visualization
 - Matplotlib Styles
 - Panda Series Plotting
 - Panda DataFrame Plotting

6. Getting Data and Manipulation

- Importing Data: csv, xml, html, xls
- APIs, web scraping
 - JSON data
 - How to read data directly from a URL
- Database
- Basic SQL

7. Web Scraping using Python

- Basic Concepts on Web Structure
- How to Download HTML Source
- Looking for Contents in HTML as XML
- Working with Web Scraping Packages (Urllib, Requests, Beautiful Soup, Selenium, Scrapy)

8. Introduction to Machine Learning with Scikit-learn

- Basic Concepts of Machine Learning
- Supervised Learning Algorithms (linear regression)
- Unsupervised Learning Algorithms (clustering)

9. Using GenAI Tool on Coding

- Demo on Text Prompt to Code Output
- Demo on Code Prompt for Debugging